Automatic Generation of Programming Exercises and Code Explanations Using Large Language Models

Year: 2022 | Citations: 416 | Authors: Sami Sarsa, Paul Denny, Arto Hellas, Juho Leinonen

---

Abstract

This article explores the natural language generation capabilities of large language models with application to the production of two types of learning resources common in programming courses. Using OpenAI Codex as the large language model, we create programming exercises (including sample solutions and test cases) and code explanations, assessing these qualitatively and quantitatively. Our results suggest that the majority of the automatically generated content is both novel and sensible, and in some cases ready to use as is. When creating exercises we find that it is remarkably easy to influence both the programming concepts and the contextual themes they contain, simply by supplying keywords as input to the model. Our analysis suggests that there is significant value in massive generative machine learning models as a tool for instructors, although there remains a need for some oversight to ensure the quality of the generated content before it is delivered to students. We further discuss the implications of OpenAI Codex and similar tools for introductory programming education and highlight future research streams that have the potential to improve the quality of the educational experience for both teachers and students alike.